Aus der Nervenklinik der Universität Tübingen
Geschäftsführender Direktor: Professor Dr. H. Heimann

Eine Untersuchung zum Vergleich zweier Einschätzskalen bei einer pharmakopsychiatrischen Studie mit akut schizophrenen Patienten

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Medizin

der Medizinischen Fakultät
(Klinische Medizin)
der Eberhard-Karls-Universität
Tübingen

vorgelegt von
HOLGER JUNG
aus Tuttlingen

1983

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41543903_0126

Aus der Nervenklinik der Universität Tübingen
Geschäftsführender Direktor: Professor Dr. H. Heimann

Eine Untersuchung zum Vergleich zweier Einschätzskalen bei einer pharmakopsychiatrischen Studie mit akut schizophrenen Patienten

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Medizin

der Medizinischen Fakultät
(Klinische Medizin)
der Eberhard-Karls-Universität
Tübingen

vorgelegt von
HOLGER JUNG
aus Tuttlingen

1983

UNIV.
LEIDEN
BIBL.

Dekan: Professor Dr. W. Frommhold
1. Berichterstatter: Professor Dr. H. Heimann
2. Berichterstatter: Professor Dr. Dr. K. Mayer
3. Berichterstatter: Professor Dr. D. Revenstorf

Druck: Zeeb-Druck, Tübingen 7 (Hagelloch)

Inhaltsverzeichnis

1.	Einleitung	S. 1
2.	Fragestellung und Methodik	S. 1
2.1	Stichprobe	S. 2
3.	Die Objektivität der Meßinstrumente	S. 3
3.1	Interrater-Reliabilität von Symptomen und Syndromen der AMDP-Skala	S. 6
3.2	Interrater-Reliabilität von Symptomen und Syndromen der BPRS-Skala	S. 13
4.	Verlaufsbeschreibung der psychopathologischen Symptomatik	S. 18
4.1	Methodik	S. 18
4.2	Verlaufsbeschreibung anhand der AMDP-Skala im Vergleich Cochran-McNemar und Friedman-Dunn-Rankin für Syndrome und Symptome	S. 25
4.3	Verlaufsbeschreibung anhand der BPRS-Skala zwischen Cochran-McNemar und Friedman-Dunn-Rankin für Syndrome und Symptome	S. 28
5.	Diskussion	S. 33
	Literaturverzeichnis	S. 44

1. Einleitung

Klinisch-pharmakologische Studien haben den Zweck, die Wirksamkeit eines Medikamentes objektiv zu überprüfen. Dabei treten Probleme bezüglich der Auswahl der Meßinstrumente und der statistischen Methoden zur Überprüfung dieser Meßinstrumente auf. In dieser Arbeit sollen diese Probleme und ihre Lösungsversuche, wie sie anläßlich einer klinisch-pharmakologischen Studie mit akut-schizophrenen Patienten auftraten, beschrieben werden. Als Medikament kam dabei das Neuroleptikum Perazin zur Anwendung.

2. Fragestellung und Methodik

Kriterien für die Auswahl eines Meßinstrumentes sind unter anderem psychometrische Kriterien wie Objektivität, Reliabilität und Validität, sowie eine leichte und ökonomische Handhabung; außerdem wie bekannt und verbreitet ein solches Instrument ist. Die Brief Psychiatric Rating Scale (BPRS) von Overall und Gorham dürfte, vor allem was den englischsprachigen Raum angeht, die am meisten benutzte Skala sein und gleichzeitig eine der kürzesten. Mit der von der Arbeitsgemeinschaft für medizinische Dokumentation entwickelten AMDP-Skala sollte andererseits der Versuch gemacht werden, ein Instrument zu entwickeln, das für den deutschsprachigen Raum, was die Verbreitung anbelangt, eine ähnliche Rolle spielen soll.

Diese beiden Skalen sollen in Bezug auf ihre Objektivität und auf ihre Brauchbarkeit als Verlaufsinstrument geprüft werden. Um eine möglichst standardisierte Untersuchungs- und Meßsituation zu gewährleisten, wurde ein festes Untersuchungsprogramm entwickelt, dem alle Patienten, die in die Studie eingingen, unterzogen wurden. Dieses Programm wurde in seinen für diese Arbeit relevanten Punkten jeweils einmal vor Beginn der Therapie (T0), einmal am 14. Tag (T14) und am 28. Tag (T28) der Therapie durchgeführt, so daß also über 3 Meßpunkte hin der Verlauf beobachtet werden konnte.
Die konkrete Untersuchungssituation sah so aus, daß von den Untersuchern ein ca. halbstündiges, halbstandardisiertes Interview durchgeführt wurde, das auf Video aufgenommen wurde. Außer psychophysiologischen und pharmakologischen Untersuchungen wurden die Patienten noch einer Untersuchung durch den Stationsarzt im Hinblick auf den Nebenwirkungsbogen der AMDP-Skala unterzogen. Die so hergestellten Videoaufnahmen wurden dann von den beiden an der AMDP-Skala trainierten Beurteilern nach der BPRS und dem psychopathologischen Befundbogen der AMDP-Skala eingeschätzt.

2.1. Stichprobe

Von den in die Studie eingegangenen 28 Patienten waren 16 männlichen und 12 weiblichen Geschlechts. Das Durchschnittsalter be-

trug für die Frauen 36,7, für die Männer 28,7 Jahre. Die Variationsbreite reichte dabei bei den Männern von 17 bis 44, bei den Frauen von 20 bis 53 Jahren (Tabelle 1).

Tabelle 1:

Pat. Nr.	Geschlecht	Alter	ICD
1	♀	39	298.9
2	♂	26	295.7
3	♀	43	296.8
4	♀	28	295.6
5	♀	53	295.3
6	♀	37	295.6
7	♀	43	295.3
8	♂	35	295.3
9	♂	21	295.3
10	♂	22	295.3
11	♂	31	295.0
12	♂	44	295.4
13	♂	26	295.0
14	♂	17	295.3
15	♀	40	295.3
16	♂	19	295.1
17	♂	21	295.1
18	♂	35	295.3
19	♀	27	295.3
20	♀	20	295.3
21	♀	38	295.3
22	♀	25	295.3
23	♂	40	298.3
24	♂	29	295.1
25	♀	48	295.4
26	♂	29	295.3
27	♂	25	298.3
28	♂	21	295.1

3. Die Objektivität der Meßinstrumente

Lienert (1969) definiert die Objektivität folgendermaßen: "Unter Objektivität eines Tests verstehen wir den Grad, in dem die Ergebnisse eines Tests unabhängig vom Untersucher sind."

Objektivität kann aufgeschlüsselt werden in Objektivität der Durchführung, der Auswertung und der Interpretation. In der vorliegenden Arbeit handelt es sich um ein Rating, d.h. um die Objektivität der Auswertung, von daher dürfte der Ausdruck Interrater-Reliabilität den Sachverhalt am genauesten beschreiben.

Wie bereits oben erwähnt, wurde die Videoaufnahme eines ca. halbstündigen, halbstandardisierten Interviews von zwei an der AMDP-Skala trainierten Beurteilern auf der AMDP und der BPRS eingeschätzt. Von dieser Anzahl der Beurteiler und der Zahl der eingehenden Personen, vom Skalenniveau sowie von den normalverteilten, bzw. nicht normalverteilten Merkmalen hängt die Wahl der zur Auswertung geeigneten Verfahren ab.

Für den hier vorliegenden Fall, zwei Beurteiler, nicht normalverteilte, polychotome Daten, ist vor allem der von Cohen entwickelte Kappakoeffizient geeignet. Cohen ging bei der Konstruktion dieses Koeffizienten von der Überlegung aus, daß ein Koeffizient, wie z.B. Pearsons Korrelationskoeffizient, der nur die allgemeine Übereinstimmung mißt, als Reliabilitätsmaß ungeeignet sei, weil hier die spezielle Urteilsübereinstimmung entscheidend sei. Verdeutlichen läßt sich das, wenn man sich eine Mehrfeldertafel vorstellt, in deren Zeilen die Schätzwerte des einen Beurteilers und in deren Spalten diejenigen des anderen Beurteilers eingetragen werden. Eine 100%ige Übereinstimmung fände sich dann, wenn nur die Diagnonallinie besetzt wäre. "Das Prinzip, nach dem Cohen einen kappa-Reliabilitäts-

koeffizienten für Kategorialdaten definiert hat, basiert auf der Beantwortung der Frage, wie weit die beobachtete Besetzungszahl der Diagonalfelder die unter der Nullhypothese unabhängiger Zeilen und Spalten erwarteten Besetzungszahlen übersteigen" (Lienert, 1978).

In dieser Arbeit soll eine spezielle Form des Kappa verwendet werden, das sogenannte "gewichtete Kappa". Diese Form des Kappa ist ursprünglich für intervallskalierte Merkmale entwickelt worden, kann aber auch für ordinal skalierte Merkmale verwendet werden, wenn eine entsprechende Gewichtung vorgenommen wird (vergl. Cicchetti, 1976). Das Prinzip ist dabei dasselbe wie beim Nominalskalen-Kappa, das oben vorgestellt wurde. Beobachtungs- und Erwartungswerte werden beide mit entsprechenden Gewichtungsfaktoren multipliziert; bei voll übereinstimmendem Rating sind die Gewichte gleich Null, während sie bei maximaler Nichtübereinstimmung ihren jeweiligen Höchstwert erreichen. Die Summe der Beobachtungswerte multipliziert mit den Gewichten, dividiert durch die Erwartungswerte multipliziert mit den Gewichten, das Ergebnis subtrahiert von 1 ergibt dann die Kappa-Reliabilität.

Daraus ergibt sich das Problem der Interpretierbarkeit von Kappa, wenn, bei der Mehrfeldertafel, Randsummen mit Null erscheinen, was bei geringer Varianz des Merkmals auftreten kann. In diesem Fall soll hier wie bei Woggon et al (1978) vorgegangen werden, indem Kappa mit der prozentualen Übereinstimmung in Bezug gesetzt wird; bei hohem Übereinstimmungs-

prozent und niedrigem Kappa ist Kappa nicht interpretierbar.

3.1 Interrater-Reliabilität von Symptomen und Syndromen der AMDP-Skala

Mombour (1977) definiert das Syndrom als "das überzufällig häufige, gemeinsame Auftreten einer Reihe von Einzelsymptomen." Es wurden bisher mehrere Untersuchungen durchgeführt, in denen durch Faktorenanalyse die Syndromstruktur der AMDP-Skala dargestellt werden sollte. Es ergaben sich dabei gute Übereinstimmungen der verschiedenen Faktorenlösungen (vergl. Sulz-Blume et al, 1979). Trotzdem wird die Faktorenstruktur wesentlich durch die Zusammensetzung der Stichprobe bestimmt. Da, bedingt durch die relativ geringe Anzahl von in diese Studie eingehenden Patienten, keine Faktorenanalyse möglich ist, soll zur Syndrombildung eine Faktorenlösung gewählt werden, die von der Stichprobenzusammensetzung der in dieser Studie eingehenden sehr ähnlich ist. Die dabei in Frage kommende ist die von Baumann (1974) durchgeführte. Danach ergeben sich 9 Faktoren, bzw. Syndrome, für die ebenfalls die Interrater-Reliabilität berechnet werden soll. Diese 9 Syndrome lauten: 1. Paranoides S. 2. Halluzinatorisches S. 3. Katatones S. 4. Apathisches S. 5. Psychoorganisches S. 6. Manisch-schizoaffektives S. 7. Hostilitäts-S.

8. Somatisch-depressives S. 9. Gehemmt-depressives S.
Nicht berücksichtigt wurde das somatisch-vegetative Syndrom, da ja die beiden Beurteiler die eingehenden Patienten nur über Videoaufnahmen einschätzen sollten und somit die dafür notwendige Untersuchung nur von den Interviewern durchgeführt werden konnte; und ebenfalls das hypochondrisch-depressive Syndrom, da es in dieser Stichprobe zu gering vorhanden war.
Für die Skalierung wurden die Kategorien "nicht beurteilbar", "fraglich vorhanden" und "nicht vorhanden" zusammengefaßt und mit den Kategorien "leicht/mittel/schwer vorhanden" von 0 bis 3 entsprechend bewertet. Für die Beurteilung der Interrater-Reliabilität soll hier nur der Zeitpunkt T0 herangezogen werden, um dem Umstand Rechnung zu tragen, daß Kappa bei abnehmender Varianz nur noch schwer zu interpretieren ist.
Betrachtet man nun in Tabelle 2 die für den Zeitpunkt T0 errechneten Kappawerte, so fallen die doch recht gleichmäßig hohen Kappawerte mit Kappa - 0,6 auf. Wenn auch diese Werte nicht unmittelbar vergleichbar sind mit den Kappawerten aus anderen vorhergegangenen Studien, sei es, weil eine andere Form des Kappa benutzt wurde, oder weil die Stichprobenzusammensetzung unterschiedlich ist, so kann man sich doch an die von Woggon (1978) vorgeschlagene Bewertung halten, nach der Kappawerte von mehr als 0,6 als gute Übereinstimmung zu betrachten sind.

Tabelle 2: Interrater-Reliabilität der AMDP-Syndrome

Syndrome	Kappa			Übereinstimmungsprozent		
	T0	T14	T28	T0	T14	T28
Para S.	0,76	0,73	0,83	71%	57%	78%
Halluz. S.	0,77	0,55	0,81	71%	53%	78%
Kataton. S.	0,63	0,43	0,8	53%	42%	82%
Apath. S.	0,92	0,83	0,79	92%	78%	75%
Psych.org. S.	0,69	0,53	0,62	67%	60%	64%
Man.-schiz. S.	0,81	0,76	0,79	78%	67%	82%
Host. S.	0,76	0,68	0,79	71%	71%	75%
Som. depr. S.	0,7	0,61	0,55	67%	60%	42%
Geh. depr. S.	0,61	0,53	0,6	53%	50%	57%

Tabelle 3: Mittelwerte und Standardabweichung der ADMP-Syndrome des 1. Beurteilers

Syndrome	Mittelwerte			Standardabweichung		
	T0	T14	T28	T0	T14	T28
Para S.	10,8	6,0	4,0	5,7	5,2	4,0
Halluz. S.	6,5	3,2	1,9	4,6	3,1	2,7
Kataton. S.	3,3	2,6	1,4	2,7	2,7	1,7
Apath. S.	2,2	1,9	1,7	2,6	2,2	1,6
Psych.org. S.	2,1	1,8	1,4	1,9	1,6	1,3
Man.-schiz. S.	3,8	2,0	1,3	4,2	3,1	3,2
Host. S.	7,7	5,5	3,5	5	4	3,5

Fortsetzung Tabelle 3:

Syndrome	Mittelwerte			Standardabweichung		
	T0	T14	T28	T0	T14	T28
Som.depr. S.	2,8	2,4	1,7	3,4	2,3	2,6
Geh.depr. S.	2,7	2,4	1,9	1,8	1,6	1,7

Tabelle 4: Mittelwerte und Standardabweichung der AMDP-Syndrome des 2. Beurteilers

Syndrome	Mittelwerte			Standardabweichung		
	T0	T14	T28	T0	T14	T28
Para S.	12,4	6,8	4,2	6,1	5,6	4,1
Halluz. S.	6,4	3,1	1,6	4,5	2,7	2,1
Kataton. S.	3,3	2,2	1,2	2,4	1,8	1,3
Apath. S.	2,1	1,8	1,7	2,5	1,8	1,8
Psych.org. S.	2,4	1,7	1,5	2,2	1,8	1,8
Man.schiz. S.	4,5	2,7	1,7	4,4	3,8	3,8
Host. S.	7,2	5,1	3,5	4,6	4,0	3,3
Som. depr. S.	3,7	2,7	1,9	3,3	2,8	1,9
Geh. depr. S.	2,8	2,3	1,8	1,9	1,8	1,9

Wichtig erscheint jedoch vor allem, daß die beiden Syndrome, die in dieser Stichprobe am höchsten laden, nämlich das paranoide und das Hostilitätssyndrom, mit einem Kappa von 0,75 eine sehr gute Übereinstimmung aufweisen. Zu einem ähnlichen Urteil kann man auch bei Betrachtung der Kappawerte der AMDP-Symptome

gelangen. Nur zwei Items, nämlich affektstarr und maniriert, liegen von den interpretierbaren unter 0,5. Interessant erscheint, daß diejenigen Items, die in der Berliner Taxilanstudie und ebenso in der in Zürich durchgeführten Studie durch ihre niedrigen Kappawerte auffielen, in dieser Studie, außer dem Item "Störung der Vitalgefühle", welches praktisch keine Varianz aufweist, sämtliche eine gute Obereinstimmung zeigen. Erwähnenswert ist weiterhin, daß von den 123 Items nur 62 in dieser Studie einen Mittelwert über 0,1 aufweisen. Dies ist sicher darauf zurückzuführen, daß die eingegangene Stichprobe relativ homogen ist. Hinweise darauf ergaben sich ja schon aus dem relativ niedrigem Score der meisten Syndrome. Auch bei Durchsicht der Mittelwerte der Items fällt auf, daß fast nur diejenigen Items, die eine Wahnsymptomatik abfragen, hohe Werte aufweisen. Oberhaupt zeigen nur 15 Items einen Mittelwert über 1. Dies ist sicher nicht daraufhin zurückzuführen, daß die in die Studie eingegangenen Patienten eine geringe Symptomatik aufwiesen, sondern darauf, daß nur akut schizophrene Patienten eingingen; d.h., das Spektrum an Symptomen ist relativ eng. Von 123 Items waren in dieser Studie insgesamt 69 interpretierbar, d.h., Kappa korrelierte einmal mit dem Obereinstimmungsprozent, zum zweiten wies das jeweilige Item eine gewisse Varianz auf, keine der Randspalten in der Mehrfeldertafel war Null.

Tabelle 5: Kappa der Einzelitems

Nr.	Item	Kappa TO	%	Mittelw.	Standardabw.
1	Bewußts.eintrüb.	1	100	0	0
2	delirant	"	"	"	"
3	umdämmert	"	"	"	"
4	andere	"	"	"	"
5	zeitlich	"	"	"	"
6	örtlich	"	"	"	"
7	situativ	"	"	"	"
8	über d. eig. Pers.	"	"	"	"
9	andere	"	"	0,1	0,5
10	Auffassungsstör.	0,62	89	0,17	0,61
11	Konz. Stör.	0,75	75	0,75	1
12	Merkfähig. Stör.	0,64	82	0,3	0,65
13	Gedächtnisstör.	1	96	0,03	0,18
14	Konfabulationen	1	100	0	0
15	andere	"	"	"	"
16	gehemmt	0,81	85	0,5	0,96
17	gesperrt	1	100	0,1	0,6
18	verlangsamt	0,63	71	0,85	1
19	eingeengt	0,69	64	1,5	1,2
20	umständlich	0,64	75	0,7	0,9
21	perseverierend	0,87	85	0,2	0,5
22	beschleunigt/ideenfl.	0,9	92	0,2	0,5
23	Vorbeireden	0,68	92	0,7	0,8
24	inkoh./zerfahren	0,56	50	1	0,8
25	Gedankenabreißen	0,73	96	0,1	0,4
26	andere	1	96	0,01	0
27	Hypochondrie	0	97	0,01	0
28	überw. Ideen	0	92	0	0
29	Zwangsdenken	1	100	0	0
30	Zwangsimpulse	"	"	"	"
31	Zwangshandlungen	"	"	"	"
32	Phobien	"	"	"	"
33	andere	1	96	0	0
34	Wahnstimmung	0,63	92	1,3	1
35	Wahnwahrnehmung	0,94	92	1,7	1,3
36	Wahneinfall	0,51	53	1,5	1,2
37	Wahndynamik	0,63	67	1,5	1,2
38	system. Wahn	0,85	89	0,9	1
39	Schuldwahn	0,65	96	0,1	0,6
40	hypoch. Wahn	0,83	89	0,4	0,8
41	Verarmungswahn	0,79	96	0,1	0,6
42	Größenwahn	0,89	89	0,7	1,2
43	Beziehungswahn	0,73	85	0,8	1,1

Nr.	Item	Kappa TO	%	Mittelw.	Standardabw.
44	Beeintr./Verfolg.Wahn	0,84	82	1	1,2
45	Eifersuchtswahn	1	100	0	0
46	anderer Wahn	0,73	92	0,1	0,5
47	Illusionen	1	100	0	0
48	Stimmenhören	0,84	89	0,5	1
49	and.akust.Halluz.	1	100	0	0
50	opt. Halluz.	"	"	"	"
51	Körperhalluz.	0,68	85	0,5	1
52	Geruchshalluz.	0,65	96	0	0,4
53	andere Störungen	0	96	0	0
54	Stör. d. Ichident.	0,91	92	1	1,3
55	and. Entfremd.Erl.	0,76	85	0,9	1,2
56	Autismus	0,72	85	0,4	0,8
57	Gedankenausbreit.	1	100	0,1.	0,5
58	Gedankenentzug	0,65	96	0,03	0,2
59	Gedankeneingebung	0,88	89	0,4	1
60	sonst. Fremdbeeinfl.E.	0,9	96	0,4	0,9
61	andere	1	100	0	0
62	ratlos	0,68	64	1	1,2
63	gefühls./affektarm	0,65	75	0,9	0,85
64	Gefühl d. Gefühllos.	0	96	0	0
65	Stör. d. Vitalgefühle	1	100	0	0
66	depremiert/traurig	0,86	89	0,7	0,8
67	hoffn.los/verzweif.	0,21	85	0,15	0,4
68	ängstlich	0,63	67	1,2	0,9
69	gehoben/euphorisch	0,63	85	0,3	0,7
70	mürrisch/dysphor.	0,85	92	0,5	1
71	mißtrauisch/feinds.	0,68	67	1	1,1
72	gespannt	0,61	64	1,1	1,2
73	innerlich unruhig	0,59	78	0,7	1
74	klagsam	0	96	0,03	0,2
75	läppisch	0,72	92	0,1	0,3
76	ekstatisch	0,36	89	0,07	0,2
77	Insuff. Gefühl	0,6	85	0,4	0,8
78	gesteig.Selbstw.Gef.	0,65	89	0,3	0,8
79	Schuldgefühl	0,67	85	0,4	0,9
80	Verarmungsgefühl	0	96	0	0
81	ambivalent	0,76	96	0,1	0,6
82	affektiv inadäquat	0,72	78	0,75	0,9
83	affektlabil	0,71	85	0,4	0,8
84	affektinkontinent	1	100	0	0
85	affektstarr	0,47	53	0,6	1
86	andere Störungen	0	92	0,1	0,6
87	antriebsarm	0,87	85	0,9	1
88	antriebsgehemmt	0,65	96	0,07	0,4
89	stuporös	0	92	0,1	0,4
90	mutistisch	0,37	92	0,1	0,4

Nr.	Item	Kappa TO	%	Mittelw.	Standardabw.
91	antriebsgesteigert	0,71	85	0,3	0,7
92	mot. unruhig	0,64	75	0,6	0,9
93	Parakinesen/Stereotyp	0,86	92	0,5	1
94	negativistisch	0,54	85	0,2	0,6
95	maniriert	0,49	67	0,8	0,9
96	Sprachzerf./Neolog.	0,78	92	0,1	0,4
97	logorrhoisch	0,54	85	0,2	0,6
98	andere Störungen	0	92	0,03	0,2
99	abends besser	1	100	0	0
100	morgens besser	"	"	"	"
101	nächtl.Exazerbat.	"	"	"	"
102	and.Verlaufsbesond.	"	"	"	"
103	Kontakt vermindert	0,81	82	0,9	1,2
104	Kontakt vermehrt	0,68	92	0,3	0,8
105	Krankheitsgef.	0,65	78	1,2	1,3
106	Mangel a. Krankh.G.	0,85	85	1,3	1,5
107	Mangel a. Krankh.Eins	0,85	67	1,8	1,4
108	Abl. d. Behandl.	0,45	89	0,2	0,7
109	Suicidtendenzen	0,65	96	0,1	0,5
110	Suicidhandlungen	1	100	0	0
111	Selbstbesch. tend.	"	"	"	"
112	Aggr. tend.	0,1	82	0,3	0,8
113	Aggr. handl.	0	96	0,1	0,5
114	Sex. gesteigert	0,79	96	0,1	0,6
115	Sex. vermindert	0,74	92	0,3	0,8
116	Sex. Triebabw.	1	100	0	0
117	Nahrungsabl.	"	"	"	"
118	ißt n. selbst.	"	"	"	"
119	unselbst. i.Körp.pfl.	"	"	"	"
120	läßt unter sich	"	"	"	"
121	Beschäft. erschw.	0	96	0,1	0,5
122	bettlägerig	1	100	0	0
123	andere Störungen	"	"	"	"

3.2 Interrater-Reliabilität von Symptomen und Syndromen der BPRS-Skala

Wir verwendeten für diese Studie eine Form der BPRS, bei der außer den 16 ursprünglichen Items noch ein zusätzliches hinzugefügt wurde, welches aber nur im Gesamtscore mitzählt. Bei der

von Mombour et al (1975) durchgeführten Faktorenanalyse wird dieses Item nicht berücksichtigt. Diese Faktorenanalyse, die die Grundlage für die Syndrombildung in dieser Arbeit darstellt, kam zu 4 Faktoren, die auch klinischen Syndromen entsprechen, nämlich: 1. Denkstörungen 2. Zurückgezogenheit - Apathie 3. Mißtrauische Feindseligkeit 4. Ängstliche Depression. Eine noch weitere Zusammenfassung führte zu einer "schizophrenic-depressive contrast function" und einer "coping-resignation contrast function", mit denen man die Patienten hinsichtlich ihrer schizophrenen bzw. depressiven Tendenz, oder ihrer mehr aktiven bzw. resignierenden Haltung beschreiben kann. Skaliert wurden die Items entsprechend ihrem Ausprägungsgrad von "nicht vorhanden" mit 1 bis "extrem" mit 7.

Ein Blick auf die Tabelle 6 zeigt, daß auch bei der BPRS-Skala die Kappawerte bei den Syndromen sämtliche eine gute Übereinstimmung zeigen. Nur die Kontrastfunktion Schizophrenie - Depression zeigt einen Wert unter 0,6.

Auch in Tabelle 9, wo die Kappawerte der Einzelitems aufgezeichnet sind, fallen die durchgängig hohen Werte auf. Hier fallen allerdings zwei Items unter die 0,6 Grenze, nämlich die Items "Ungewöhnliche Denkinhalte" und "Verflachung, affektive Abstumpfung". Allerdings muß bei beiden bemerkt werden, daß sie gegenüber den restlichen Items etwas weniger konkret inhaltlich ausformuliert wurden, sodaß der Bewerter hier mit seinen eigenen Wertvorstellungen einen etwas weiteren Bewertungsspielraum erhält.

Insgesamt kann man hier schon festhalten, daß gegenüber der AMDP-Skala einmal sämtliche Items interpretierbar sind; zum zweiten, daß die Items, relativ zur AMDP-Skala gesehen, einen insgesamt höheren Mittelwert aufweisen. Auf mögliche Gründe dafür soll in der abschließenden Diskussion eingegangen werden.

Tabelle 6: Interrater-Reliabilität der BPRS-Syndrome und der Kontrastfunktionen

Syndrome	Kappa			Obereinstimmungsprozent		
	T0	T14	T28	T0	T14	T28
Denkstörungen	0,71	0,61	0,52	64	67	53
Apathie	0,68	0,71	0,47	64	64	57
Feindseligk.	0,82	0,73	0,7	78	67	67
Ängstl. Depr.	0,74	0,76	0,53	67	75	50
Schizo-Depr.CF	0,55	0,56	0,55	50	57	60
Cop.-Resig. CF	0,7	0,69	0,47	64	67	57

Tabelle 7: Mittelwerte und Standardabweichung der BPRS-Syndrome und Kontrastfunktionen des 1. Beurteilers

Syndrome	Mittelwerte			Standardabweichung		
	T0	T14	T28	T0	T14	T28
Denkstörungen	11,2	7,8	6,3	3	3,2	3,2
Apathie	9,6	9,6	9,2	4,3	3,9	3,9

Fortsetzung Tabelle 7:

Syndrome	Mittelwerte			Standardabweichung		
	T0	T14	T28	T0	T14	T28
Feindseligkeit	10,6	7,8	7	4,3	4,5	4,1
Ängstl. Depression	8,4	7,6	6,7	3,5	3,3	3,1
Schizo.-Depr. CF	1,9	0,2	-0,4	5,3	4,6	3,9
Coping-Resign. CF	4,1	0,2	-0,5	6,3	6	5,8

Tabelle 8: Mittelwerte und Standardabweichung der BPRS–Syndrome und Kontrastfunktionen des 2. Beurteilers

Syndrome	Mittelwerte			Standardabweichung		
	T0	T14	T28	T0	T14	T28
Denkstörungen	11,2	8,1	6,5	3,4	3,2	3,2
Apathie	9,4	8,6	8,2	3,9	3,8	3,6
Feindseligkeit	10,2	7,3	6,7	3,8	3,8	3,7
Ängstl. Depression	8,2	7,2	6,4	3,1	3,2	2,8
Schizo.-Depr. CF	3	0,8	0,1	4,8	4,5	3,9
Coping-Resign. CF	3,6	0,9	-0,4	6,5	5,7	5

Tabelle 9: Kappa der Einzelitems der BPRS-Skala

Nr.	Item	Kappa TO	%	Mittelw.	Standardabw.
1	Körperbezogenheit	0,71	57	3	1,8
2	Angst	0,63	57	3,9	1,4
3	Emot. Zurückgezogenh.	0,74	57	3,7	1,7
4	Zerfall d.Denkprozesse	0,62	42	2,9	1,4
5	Schuldgefühle	0,74	78	1,8	1,7
6	Gespanntheit	0,75	57	3,5	1,7
7	Maniriertheit	0,64	50	2,9	1,3
8	Größenerleben	0,82	75	2,2	2,1
9	Depression	0,67	67	2,5	1,5
10	Feindseligkeit	0,8	71	3,3	1,9
11	Mißtrauen	0,75	71	4,2	1,9
12	Halluzinationen	0,9	85	2,4	2
13	Antriebsminderung	0,63	57	2,8	1,7
14	mangelnde Kooperation	0,75	75	2,9	2,2
15	Ungewöhnl. Denkinh.	0,57	50	5,6	1,1
16	Verflachung	0,54	53	3,2	1,8
17	Gedankenverarmung	0,73	64	2,2	1,8

4. Verlaufsbeschreibung der psychopathologischen Symptomatik

Im zweiten Teil dieser Arbeit soll der Verlauf der 28tägigen Neuroleptikatherapie anhand der AMDP-und der BPRS-Skala beschrieben werden. Aus Gründen der besseren Übersicht soll dabei nur auf die Ergebnisse des 1. Beurteilers eingegangen werden, was angesichts der guten Interrater-Reliabilität auch naheliegt. Überdies soll auch ein Vergleich mit den Ergebnissen der Berliner Taxilanstudie angestellt werden.

4.1 Methodik

Um einen Vergleich anstellen zu können, sollte die Methodik bei zwei verschiedenen Untersuchungen möglichst gleich sein. Für die Auswertung der Berliner Taxilanstudie wurde einmal die Q-Statistik von Cochran benutzt, zum zweiten, um die signifikante Richtungsänderung festzustellen, der Symmetrietest von McNemar. Außerdem wurde als zusätzliches beschreibendes Moment eine Weiterentwicklung des Ausprägungs- und Remissionsquotienten benutzt, der Besserungs- (BQ) und Provokationsquotient (PQ). Da in dieser Arbeit die Daten nicht nur dichotomisiert werden sollten, um jeden Informationsverlust zu vermeiden, wurden zusätzlich zwei weitere Verfahren benutzt, die zueinander ähnlich stehen wie die Q-Statistik und der Symmetrietest: der Friedman-Test und der d_{ij}-Test von Dunn-Rankin. Beide bieten den einen Vorteil, daß die Daten nicht dichotomisiert werden müssen. Der Friedman-Test dient in der Hauptsache der Auswertung von sogenannten Kendall-Plänen, das heißt, N Individuen werden den k Stufen

eines Faktors, in diesem Fall eine Behandlung, ausgesetzt. Es werden also k abhängige Stichproben mit einem nicht-parametrischen Verfahren geprüft. Ein signifikantes Ergebnis besagt allerdings nur, daß zumindest zwei der geprüften Stichproben sich unterscheiden. Zur Beantwortung der Frage, welche, bzw. wieviele der verschiedenen Populationen sich unterscheiden, wird der Dunn-Rankin-Test benutzt. Dieser Test unterscheidet sich noch zusätzlich dadurch vom Symmetrietest von McNemar, daß hierbei alle Populationen gleichzeitig gegeneinander geprüft werden und nicht nur ein Paarvergleich stattfindet. Das führt dazu, daß der Dunn-Rankin-Test in Relation Cochran-McNemar und Friedman-Dunn-Rankin etwas strenger prüft als der McNemar-Test. Dies sollte jedoch ausgeglichen werden durch den zusätzlichen Informationsgewinn durch den Eingang der polychotomen Daten. Ein weiterer Unterschied besteht noch darin, daß der Dunn-Rankin-Test nicht wie der McNemar die Richtung der Veränderung durch ein Vorzeichen anzeigt. Hierfür müssen zusätzlich die Verlaufsmaße Besserungs- und Provokationsquotient herangezogen werden, sowie die Mittelwerte der Items zu den verschiedenen Zeitpunkten. Die Verlaufsmaße Besserungs- und Provokationsquotient sind dabei so definiert, daß der Besserungsquotient eine Abnahme der Häufigkeit von vorher manifesten Merkmalen beschreibt, der Provokationsquotient ein Neuauftreten von Merkmalen. Ein BQ von 1 bedeutet dann das völlige Verschwinden eines Symptoms bei allen Merkmalsträgern, während ein PQ von 1 ausdrückt, daß bei allen vorher merkmalsfreien das Symptom provo-

ziert wurde.

Zu den Möglichkeiten des Friedman und Dunn-Rankin-Tests sollte noch nachgetragen werden, daß bei diesen beiden Verfahren, dadurch, daß die Daten nicht nachträglich dichotomisiert werden, eine Verlaufsbeschreibung auch für den Gesamtscore und die Syndrome methodisch unbedenklich ist und somit durchgeführt werden soll.

Tabelle 10: Mittelwerte der einzelnen Merkmale an den Zeitpunkten T0, T14, T28 sowie das Signifikanzniveau der Q-Statistik der AMDP-Skala

Nr.	Item	Mittelwerte T0	T14	T28	%
24	inkohärent	1	0,4	0,2	0,1
34	Wahnstimmung	1,3	0,6	0,4	"
19	eingeengt	1,6	1	0,9	1
23	Vorbeireden	0,7	0,3	0,1	"
35	Wahnwahrnehmung	1,7	1,1	0,6	"
36	Wahneinfall	1,6	0,7	0,5	"
38	systemat. Wahn	0,8	0,5	0,2	"
48	Stimmenhören	0,5	0,1	0,1	"
82	affektiv inadäquat	0,8	0,4	0,3	"
95	maniriert	0,8	0,4	0,07	"
105	Krankheitsgefühl	1,2	1,4	1,3	"
37	Wahndynamik	1,5	0,9	0,6	5

Fortsetzung Tabelle 10:

Nr.	Item	Mittelwerte T0	T14	T28	%
42	Größenwahn	0,7	0,14	0,07	5
54	Störung d. Ichidentität	1	0,6	0,3	"
69	euphorisch	0,3	0,14	0,1	"
71	mißtrauisch	1	0,8	0,4	"
83	affektlabil	0,4	0,1	0,03	"
92	mot. unruhig	0,6	0,4	0,1	"
93	Parakinesen/Stereotyp.	0,5	0,1	0,03	"
97	logorrhoisch	0,3	0,14	0,07	"
106	Mangel a. Krankh.gefühl	1,2	0,8	0,5	"
112	Aggressionstend.	0,3	0	0,07	"

Tabelle 11: Signifikante Änderungen der Einzelmerkmale im Symmetrietest nach McNemar

Nr.	Item	T0/T14	T0/T28
19	eingeengt		5 %
23	Vorbeireden	5 %	1 %
24	inkohärent		1 %
34	Wahnstimmung		1 %
35	Wahnwahrnehmung		5 %
36	Wahneinfall	5 %	5 %
38	systemat. Wahn		1 %
82	affektiv inadäquat		5 %
95	maniriert		1 %
105	Krankheitsgefühl		5 %

Tabelle 12: Signifikante Änderungen der Einzelmerkmale der AMDP-Skala im Friedman-Test

Nr.	Item	Friedman-Test %
23	Vorbeireden	0,1
24	inkohärent	"
34	Wahnstimmung	"
35	Wahnwahrnehmung	"
36	Wahneinfall	"
37	Wahndynamik	"
95	maniriert	"
19	eingeengt	1
38	systemat. Wahn	"
42	Größenwahn	"
43	Beziehungswahn	"
44	Beeinträchtigungs/Verfolgungswahn	"
54	Störung der Ichidentität	"
62	ratlos	"
71	mißtrauisch/feindselig	"
82	affektiv inadäquat	"
93	Parakinesen/Stereotypien	"
51	Körperhall./Körperfühlstörungen	5
55	andere Entfremdungserlebnisse	"
60	sonstige Fremdbeeinflußungserlebnisse	"
69	gehoben/euphorisch	"
83	affektlabil	"
92	motorisch unruhig	"

Fortsetzung Tabelle 12:

Nr.	Item	Friedman-Test %
97	logorrhoisch	5
106	Mangel a. Krankheitsgefühl	"
112	Aggressionstendenzen	"

Tabelle 13: Signifikante Änderungen der Einzelmerkmale der AMDP-Skala im Dunn-Rankin-Test

Nr.	Item	Dunn-Rankin T0/T14	T0/T28
23	Vorbeireden		5 %
24	inkohärent/zerfahren	5 %	1 %
34	Wahnstimmung		1 %
35	Wahnwahrnehmung		1 %
36	Wahneinfall		1 %
37	Wahndynamik		1 %
95	maniriert		5 %

Tabelle 14: Signifikante Änderungen der Syndrome der AMDP-Skala im Friedman und Dunn-Rankin-Test sowie die Mittelwerte

	Friedman	Dunn-Rankin		Mittelwerte		
		T0/T14	T0/T28	T0	T14	T28
Gesamtscore	0,1 %	1 %	1 %	43,4	29,5	21,5
Paranoides Syn.	0,1 %	1 %	1 %	10,8	6,0	4,0
Halluz. "	0,1 %	1 %	1 %	6,5	3,2	1,9
Katatones "	1 %		5 %	3,3	2,6	1,4
man. schiz."	0,1 %	5 %	1 %	3,8	2,0	1,3

Fortsetzung Tabelle 14:

	Friedman	Dunn-Rankin		Mittelwerte		
		T0/T14	T0/T28	T0	T14	T28
Hostilitäts Syn.	0,1 %		1 %	7,7	5,6	3,5
som. depr. "	5 %			2,9	2,4	1,7

Tabelle 15: Verlaufsmaße der im Dunn-Rankin signifikant veränderten Einzelmerkmale

Nr.	Item	PQ		BQ	
		T0/T14	T0/T28	T0/T14	T0/T28
23	Vorbeireden	0,07	0,07	0,6	0,8
24	inkohärent/zerfahren	0,1	0	0,4	0,6
34	Wahnstimmung	0	0	0,3	0,5
35	Wahnwahrnehmung	0	0	0,2	0,3
36	Wahneinfall	0,1	0,1	0,5	0,6
37	Wahndynamik	0,1	0,1	0,2	0,3
95	maniriert	0,3	0,1	0,6	1

Tabelle 16: Mittelwerte der im Dunn-Rankin signifikant veränderten Einzelmerkmale

Nr.	Item	Mittelwerte		
		T0	T14	T28
23	Vorbeireden	0,7	0,3	0,1
24	inkohärent/zerfahren	1	0,4	0,2
34	Wahnstimmung	1,3	0,6	0,4
35	Wahnwahrnehmung	1,7	1,1	0,6
36	Wahneinfall	1,6	0,7	0,5
37	Wahndynamik	1,5	0,9	0,6

Fortsetzung Tabelle 16:

Nr.	Item	Mittelwerte		
		T0	T14	T28
95	maniriert	0,8	0,4	0,07

4.2 Verlaufsbeschreibung anhand der AMDP-Skala im Vergleich Cochran - McNemar und Friedman - Dunn-Rankin für Syndrome und Symptome

Von den 123 Merkmalen, die auf dem psychopathologischen Befundbogen erhoben werden, zeigten, wie aus Tabelle 10 ersichtlich ist, in der Q-Statistik 22 Items eine signifikante Veränderung, davon 2 auf dem 0,1 % Niveau, 9 auf dem 1 % Niveau, der Rest auf dem 5 % Niveau. Im Friedman-Test zeigten dagegen 26 Items eine signifikante Veränderung, 7 davon auf dem 0,1 % Niveau, 10 auf dem 1 % Niveau, der Rest auf dem 5 % Niveau. Dies kann so interpretiert werden, daß der Friedman-Test durch den Eingang von polychotomen Daten gegenüber der Q-Statistik eine größere Meßbreite besitzt.

Bemerkenswert, daß vor allem Items wie Wahnwahrnehmung, Wahneinfall, Wahndynamik, Wahnstimmung, eingeengt, also Items, die typisch sind für eine Wahnsymptomatik, sehr hohe Mittelwerte zum Zeitpunkt T0 aufweisen, und auch, vor allem im Friedman-Test, die stärkste Veränderungstendenz.

Die Items, die im Friedman-Test, bzw. in der Q-Statistik signifikante Veränderungen zeigten, wurden zur weiteren Analyse dem

Symmetrietest von McNemar, bzw. dem Dunn-Rankin-Test unterzogen, um einmal eventuell den Zeitpunkt der Veränderung genauer festlegen zu können, zum zweiten, um, wenigstens mit dem McNemar-Test, die Richtung der Veränderung festzustellen. Insgesamt zeigten 10 Items signifikante Richtungsänderungen im Symmetrietest, davon nahm eines, nämlich das Item"Krankheitsgefühl", zwischen den Zeitpunkten T0 und T28 an Häufigkeit zu, alle anderen nahmen ab. Zwei davon, nämlich "Vorbeireden" und "Wahneinfall" schon zwischen den Zeitpunkten T0 und T14, die anderen zwischen T0 und T28. Auch im Symmetrietest stellte sich eine signifikante Abnahme auf dem 1 % Niveau der Items "Wahnstimmung", "systematisierter Wahn", "Vorbeireden", dar; die Items "Wahnwahrnehmung" und "Wahneinfall" nahmen auf dem 5 % Niveau ab.
Im Dunn-Rankin-Test wiesen im Vergleich zum McNemar nur 7 Items eine signifikante Veränderung auf, davon eines, nämlich "inkohärent/zerfahren", schon zwischen den Zeitpunkten T0 und T14, alle anderen zwischen T0 und T28. Auffallend war auch hier wieder die Veränderung auf dem 1 % Niveau der Items "Wahnstimmung", "Wahneinfall", "Wahnwahrnehmung" und "Wahndynamik". Es läßt sich im direkten Vergleich zwischen McNemar und Dunn-Rankin also sagen, daß der Dunn-Rankin, obwohl mehr Information in ihn eingeht, offensichtlich wesentlich enger mißt; dies könnte darin begründet sein, daß er nicht nur wie der McNemar einen Paarvergleich anstellt, sondern daß hier alle Meßpunkte gleichzeitig in den Test eingehen. In beiden Tests zeigen diejenigen Items, die eine Wahnsymptomatik abfragen, die beste Rückbildungstendenz. Dies schlägt sich auch

in den Verlaufsmaßen nieder, wo der Provokationsquotient bei den Items "Wahnstimmung" und "Wahnwahrnehmung" Null ist, bei den Items "Wahneinfall" und "Wahndynamik" mit 0,1 nur minimal erhöht, der Besserungsquotient vor allem bei den Items "Wahneinfall" und "Wahnstimmung" eine deutlich rückläufige Tendenz aufweist. Nicht erklärbar ist, daß das Item "Wahndynamik", welches im Dunn-Rankin zwischen T0 und T28 deutlich gebessert erscheint, sich im McNemar nicht signifikant verändert, obwohl es, wie ein Blick auf Tabelle 16 zeigt, bei den Mittelwerten ebenfalls eine deutliche Besserung vorhanden ist. Im Vergleich mit der Berliner Taxilan-Studie läßt sich bisher also festhalten, daß die Veränderungen sich in dieser Stichprobe erst wesentlich später manifestieren, und zwar im McNemar nur bei zwei Items bereits am 14. Tag, im Dunn-Rankin nur bei einem, den Items "Vorbeireden", "Wahneinfall" und "inkohärent/zerfahren", von denen sich in der Berliner Studie die beiden letzteren bereits am 5. Behandlungstag signifikant änderten, während das Item "Vorbeireden" sich im McNemar überhaupt nicht signifikant veränderte. Es war allerdings auch bei der Berliner Studie auffallend, daß sich hochsignifikante Änderungen auch erst am 30. Behandlungstag zeigten.

Insgesamt zeigten sich in der Berliner Studie 62 Items in der Q-Statistik signifikant verändert, davon erschienen 16 auch in dieser Studie in der Q-Statistik, die zusätzlich signifikanten Items sind: "maniriert", "Krankheitsgefühl", "Größenwahn", "gehoben/euphorisch", "Parakinesen/Stereotypien","Aggressionstendenzen".

Von den 10 Items, die in dieser Studie im McNemar eine signifikante Richtungsänderung aufwiesen, waren nur 4 im McNemar der Berliner Studie verändert. Diese gemeinsamen Items waren: "Wahnstimmung", "inkohärent/zerfahren", "Wahneinfall/Wahngedanken", "Wahnwahrnehmungen".
Eine Verlaufsbeschreibung des Gesamtscore und der Syndrome kann mit der Q-Statistik und dem McNemar aus methodischen Gründen nicht geleistet werden, von daher ist der Vergleich mit der Berliner Studie nicht möglich. Ein Blick auf Tabelle 14 zeigt für den Gesamtscore eine signifikante Rückbildungstendenz schon am 14. Tag, ebenso für das Paranoide und das Halluzinatorische Syndrom, welche auch, zusammen mit dem Hostilitäts Syndrom die höchsten Faktorenladungen überhaupt aufweisen. Auch dies ist ein weiterer Hinweis auf die Homogenität der Stichprobe.

4.3 Verlaufsbeschreibung anhand der BPRS-Skala im Vergleich zwischen Cochran - McNemar und Friedman - Dunn-Rankin für Symptome und Syndrome

Von den 17 Items der BPRS-Skala zeigen sich, wie aus Tabelle 19 ersichtlich ist, 5 in der Q-Statistik signifikant verändert, davon 2 auf dem 0,1 % Niveau, eines auf dem 1 % Niveau und zwei auf dem 5 % Niveau. Im Friedman-Test dagegen erscheinen 10 Items verändert, zwei auf dem 0,1 % Niveau, 5 bei 1 % und 3 bei 5 %. Auch hier zeigt sich wieder, wie schon bei der AMDP-Skala, daß der

Friedman-Test mehr Informationen gibt als die Q-Statistik. Es fällt auch bei der BPRS-Skala auf, daß vor allem Items, die Denk- und Affektstörungen abfragen, signifikant verändert sind. Die Items, die in der Q-Statistik und im Friedman-Test signifikant erschienen, sollen auch hier wieder einer weiteren Analyse durch den McNemar-, bzw. Dunn-Rankin-Test zugeführt werden. Im Symmetrietest erschienen insgesamt 3 Items verändert, dabei ist interessant, daß das Item "Zerfall der Denkprozesse" bereits am 14. Tag eine deutliche Besserung aufweist, die bis zum 28. Tag noch weiter zunimmt. Dagegen zeigt das Item "Mißtrauen/ paranoide Inhalte" eine Besserungstendenz bis zum 14. Tag, die dann jedoch wieder abnimmt. Dieser Trend drückt sich auch in den Mittelwerten und in den Verlaufsmaßen aus, jedoch nicht im Dunn-Rankin-Test. Zu den signifikanten Items im Symmetrietest kommen im Dunn-Rankin noch zwei weitere Items hinzu, nämlich "Angst" und "ungewöhnliche Denkinhalte". Dabei verbessert sich "Angst" ebenfalls erst am 28. Tag, "ungewöhnliche Denkinhalte" bereits am 14. Tag und dieser Trend hält weiter an bis zum 28. Tag. Insgesamt nahmen alle Items, die im McNemar und Im Dunn-Rankin signifikant erschienen, in ihren Häufigkeiten ab. Dies drückt sich auch in den Verlaufsmaßen aus. Außer dem Item "Maniriertheit" sind bei allen Items, die im Dunn-Rankin signifikant verändert erscheinen, die Provokationsquotienten Null, während die Besserungsquotienten eine eindeutige Zunahme aufweisen, am deutlichsten bei dem Item "Zerfall der Denkprozesse" und am schwächsten bei den Items "Angst" und "ungewöhnliche Denkinhalte". Bei dem Item "Mißtrauen/paranoide Inhalte" erscheint

Tabelle 17: Signifikante Änderungen der Syndrome und des Gesamtscore der BPRS-Skala und ihre Mittelwerte

	Friedman	Dunn-Rankin		Mittelwerte		
		T0/T14	T0/T28	T0	T14	T28
Gesamtscore	0,1 %	1 %	1 %	51,5	42,0	37
Denkstörungen	0,1 %	1 %	5 %	11,2	7,8	6,3
Mißtrauische Feinds.	0,1 %	1 %	1 %	10,6	7,8	7,0
Ängstliche Depression	0,1 %		1 %	8,4	7,6	6,7
Funktion S - D	1 %		1 %	1,9	0,2	-0,4
Funktion C - R	0,1 %	1 %	1 %	4,1	0,2	-0,6

Tabelle 18: Signifikante Änderungen der Einzelmerkmale der BPRS-Skala und ihre Mittelwerte

Nr.	Item	Friedman	Dunn-Rankin		Mittelwerte		
			T0/T14	T0/T28	T0	T14	T28
2	Angst	1 %		5 %	3,9	3,3	3,0
4	Zerf. d. Denkpr.	1 %	1 %	1 %	2,9	1,9	1,6
6	Gespanntheit	5 %			3,5	2,8	2,5
7	Maniriertheit	0,1 %		1 %	2,9	2,0	1,4
8	Größenerleben	5 %			2,2	1,4	1,3
10	Feindseligkeit	1 %			3,3	2,5	2,4
11	Mißtrauen/paran.I	0,1 %	5 %	1 %	4,2	3,1	2,8
12	Halluzinationen	5 %			2,4	1,7	1,5
14	Mangelnde Koop.	1 %			2,9	2,1	1,7
15	Ungewöhnl.Denkinh	0,1 %	1 %	1 %	5,6	4,2	3,4

Tabelle 19: Signifikante Veränderungen der Einzelmerkmale der BPRS-Skala in der Q-Statistik und im Symmetrietest

Nr.	Item	Cochran	McNemar	
			T0/T14	T0/T28
4	Zerfall d. Denkprozesse	0,1 %	1 %	0,1 %
7	Maniriertheit	0,1 %		1 %
10	Feindseligkeit	5 %		
11	Mißtrauen/paran.Inh.	1 %	5 %	
12	Halluzinationen	5 %		

Tabelle 20: Verlaufsmaße der signifikant veränderten Items der BPRS-Skala im Dunn-Rankin-Test

Nr.	Item	BQ		PQ	
		T0/T14	T0/T28	T0/T14	T0/T28
2	Angst	0,07	0,07	0	0
4	Zerfall d. Denkprozesse	0,5	0,5	0	0
7	Maniriertheit	0,3	0,6	0,1	0,1
11	Mißtrauen/paran. Inhalte	0,3	0,2	0	0
15	Ungewöhnl. Denkinhalte	0,1	0,1	0	0

der Besserungsquotient zwischen T0 und T28 niedriger als zwischen T0 und T14, während er bei dem Item "Zerfall der Denkprozesse"zwischen T0 und T14 und T0 und T28 fast gleich hoch ist, ebenso wie bei den Items "Angst" und "ungewöhnliche Denkinhalte".

Für den Gesamtscore und die Faktoren läßt sich im Friedman und Dunn-Rankin-Test eine ähnlich gute Rückbildungstendenz feststellen wie bei der AMDP-Skala. Der Gesamtscore zeigt sich im Friedman-Test auf dem 0,1 % Niveau verändert, beim Dunn-Rankin zwischen T0 und T14 ebenso zwischen T0 und T28 auf dem 1 % Niveau, das gleiche gilt für den Faktor "Mißtrauische Feindseligkeit".

Bei dem Faktor "Denkstörungen" zeigt sich eine Besserung interessanterweise vor allem zwischen T0 und T14, weniger stark zwischen T0 und T28, was sich auch in den Mittelwerten ausdrückt. Bei dem Faktor "ängstliche Depression" tritt die Besserung dagegen erst in der zweiten Hälfte der Behandlung auf, was sich bei den Mittelwerten ebenfalls nachvollziehen läßt.

Bezüglich der Verlaufsbeschreibung anhand der Funktionsformeln läßt sich feststellen, daß bei der Funktion schizophren versus depressiv ein Umschwung zugunsten der depressiven Symptomatik erst in der zweiten Hälfte stattfindet, während bei der Funktion, die die mehr aktive versus die mehr resignative Einstellung beschreibt, ein solcher Umschwung sich schon am 14. Tag beobachten läßt, welcher auch noch bis zum 28. Tag anhält.

Insgesamt lassen sich die Ergebnisse der Verlaufsbeschreibung anhand der BPRS-Skala so zusammenfassen, daß auch hier wie bei der

AMDP-Skala sich eine Besserung auf Itemniveau zum großen Teil erst am 28. Tag abzeichnet, während die Besserung auf Syndromebene schon ab dem 14. Tag deutlich wird, zumindest bei jenen Faktoren, die mehr eine Wahnsymptomatik abfragen. Ein ähnlicher Trend besteht auch auf Itemebene bei jenen Items, die paranoide, bzw. wahnhafte Inhalte abfragen. Auch dies ist wieder ein deutlicher Hinweis auf die Homogenität der Stichprobe.

5. Diskussion

In dieser Arbeit sollte ein Vergleich der "Brief Psychiatric Rating Scale" (BPRS) von Overall und Gorham mit dem von der "Arbeitsgemeinschaft für Methodik und Dokumentation in der Psychiatrie" (AMDP) herausgegebenen AMDP-System bezüglich der Interrater-Reliabilität und der Brauchbarkeit als Instrument zur Verlaufsbeschreibung erfolgen. Es sollen zunächst die Ergebnisse der Interrater-Reliabilität nochmals kurz zusammengefaßt und diskutiert werden.
Von den 123 Items des psychopathologischen Befundbogens der AMDP-Skala sind bei 69 Items die Kappakoeffizienten interpretierbar. Davon liegen 11 Items unter einem Kappa von 0,6, nämlich die Items: "inkohärent /zerfahren", "Wahneinfall/Wahngedanken", "innerlich unruhig", "ekstatisch", "affektstarr", "mutistisch", "negativistisch", "maniriert", "logorrhoisch", "Mangel an Krankheitseinsicht", "Ablehnung der Behandlung". Drei davon, nämlich"affektstarr", "maniriert", "mutistisch" liegen unter einem Kappa von 0,5. Insgesamt

haben von den 123 Items 62 einen Mittelwert über 0,1, davon 15 einen über 1. Alle Items, außer dem Item "ekstatisch", die ein Kappa unter 0,6 haben, weisen auch einen Mittelwert über 0,1 auf; die Items "inkohärent/zerfahren", "Wahneinfall/Wahngedanken" und "Mangel an Krankheitseinsicht" sogar über 1.

Wie bereits oben erwähnt, wurden auch für übergeordnete Syndrome Kappawerte berechnet. Dabei lagen für alle 9 Syndrome die Kappawerte über 0,6. Die niedrigsten Werte wiesen das "gehemmt depressive Syndrom" mit 0,61 und das "katatone Syndrom" mit 0,63 auf. Vor allem das "paranoide", das "halluzinatorische" und das "Hostilitätssyndrom" wiesen mit Kappakoeffizienten von 0,76, 0,77 und 0,76 sehr gute Werte auf; gleichzeitig waren es auch diejenigen Syndrome, die mit Werten von 10,8, 6,5 und 7,7 die mit Abstand höchsten Mittelwerte aufwiesen.

Bei der BPRS-Skala fiel einmal auf, daß sämtliche Items bezüglich der Kappawerte interpretierbar waren. Zum zweiten, daß nur die Items "ungewöhnliche Denkinhalte" mit 0,57 und "Verflachung" mit 0,54 unter der 0,6 Grenze lagen. Die Items "ungewöhnliche Denkinhalte" und "Mißtrauen/paranoide Inhalte" wiesen mit 5,6 und 4,2 die höchsten Mittelwerte auf. Im übrigen lagen die Mittelwerte relativ gleichmäßig zwischen 1,8 ("Schuldgefühle") und 3,9 ("Angst").

Bei den Syndromen der BPRS inklusive der Kontrastfunktionen fielen die Kappawerte, außer bei der Kontrastfunktion "Schizophrenie versus Depression" mit 0,55, alle über 0,6 aus. Der Faktor mit dem niedrigsten Wert war "Apathie" mit 0,68. Die Faktoren "Denkstörungen" und"Feindseligkeit" zeigten mit Werten von 11,2 und 10,2 die höchsten Mittel-

werte. Die Mittelwerte der Kontrastfunktionen können nur in dem Sinn bewertet werden, als ein Umschlagen von positiven zu negativen Werten, bzw. umgekehrt, ein Zeichen für das Überwiegen der depressiven oder schizophrenen, bzw. resignativen oder mehr aktiven Komponente darstellt.

Zusammenfassend läßt sich zur Interrater-Reliabilität der AMDP-und der BPRS-Skala sagen, daß die Kappawerte bei der BPRS für Symptome und Syndrome insgesamt gut waren. Nur zwei Symptome lagen etwas unter dem Durchschnitt, was schon in der Arbeit von Mombour et al (1975) aufgefallen war. Bei der AMDP war auffällig, daß ein großer Teil der Items bezüglich des Kappas nicht zu interpretieren war. Dies dürfte daran liegen, daß die hier untersuchte Stichprobe sehr homogen war und von daher ein Großteil der Merkmale, die von den entsprechenden Items abgefragt werden, keinerlei Varianz aufwiesen. Dies verdeutlichte sich auch in der geringen Anzahl von Items, die hohe Mittelwerte zeigten. Es kann auf der anderen Seite auch als eine Bestätigung der Annahme von Freudenthal, Gebhardt und Pietzcker (1977) interpretiert werden, die ja die Items der BPRS eher als Syndrome im Sinne derjenigen, durch Faktorenanalyse aus der AMDP gewonnenen, auffaßten. Insgesamt scheint die BPRS aber als Meßinstrument, wenigstens was ihre Interrater-Reliabilität betrifft, vor allem bei homogenen Stichproben, in ihrer derzeitigen Form gegenüber der AMDP günstiger zu sein.

Die Methode zur Analyse des Verlaufs wurde weiter oben schon ausführlich besprochen. Es sollen an dieser Stelle noch einmal kurz die Ergebnisse beider Auswertungsmethoden und vor allem die aufge-

tretenen Unterschiede dargestellt werden.

In der Q-Statistik der AMDP erschienen 22 Items signifikant verändert gegenüber 26 im Friedman-Test. Dabei waren 2 gegenüber 7 auf dem 0,1 % Niveau verändert, 9 gegenüber 10 auf dem 1 % Niveau und 9 gegenüber 9 auf dem 5 % Niveau. Die Items, die im Friedman-Test hinzukamen, waren: "Beziehungswahn", "Beeinträchtigungs-/Verfolgungswahn", "ratlos", (alle 1 %); "Körperhalluzinationen/Körpergefühlsstörungen", "andere Entfremdungserlebnisse", "sonstige Fremdbeeinflussungserlebnisse", "Aggressionstendenzen", (alle 5 %). Im Symmetrietest nach McNemar erschienen bei der AMDP-Skala mehr Items signifikant verändert als im Dunn-Rankin-Test; und zwar erschienen im Dunn-Rankin nur diejenigen Items, die im Friedman-Test auf dem 0,1 % Niveau signifikant verändert waren, während im Symmetrietest auch Items erschienen, die sich in der Q-Statistik auf dem 1 % Niveau verändert zeigten. Im Symmetrietest nach McNemar waren folgende Items zusätzlich signifikant: "eingeengt", "Wahndynamik", "systematisierter Wahn", "affektiv inadäquat", (alle zwischen T0 und T28 abgenommen) "Krankheitsgefühl" (zwischen T0 und T28 zugenommen). Dabei veränderten sich im Symmetrietest zwei Items bereits am 14. Tag, nämlich "Vorbeireden" und "Wahneinfall" (beide auf dem 5 % Niveau), im Dunn-Rankin-Test nur eines: "inkohärent/zerfahren" (ebenfalls auf dem 5 % Niveau), alle anderen zeigten eine signifikante Änderung erst am 28. Tag. Diese Änderung der drei Items am 14. Tag läßt sich auch anhand der Mittelwerte nachvollziehen, die ebenfalls zwischen T0 und T14 stärker fallen als in der zweiten Hälfte der Behandlung. Prinzipiell läßt sich zu den Mittelwerten der Items der AMDP-Skala sagen, daß hauptsächlich die-

jenigen Items einen relativ hohen Mittelwert aufweisen, welche eine Wahnsymptomatik abfragen. Wie bereits weiter oben erwähnt, weisen nur 15 Items einen Mittelwert über 1 auf (bei einer Skalierung zwischen 0 und 3), die davon im Symmetrietest signifikant veränderten Items sind (nach der Höhe ihrer Mittelwerte): "Wahnwahrnehmung", "eingeengt", "Wahneinfall", "Wahnstimmung", "Krankheitsgefühl", "inkohärent/zerfahren".

Von den im Symmetrietest veränderten Items wies in den Verlaufsmaßen nur das Item "maniriert" eine gewisse Provokationstendenz zwischen T0 und T14 auf, alle anderen hatten entweder nur eine minimale Erhöhung des Provokationsquotienten oder einen von Null. Beim Besserungsquotienten wiesen alle Items eine zunehmende Besserung auf, am stärksten die Items "maniriert" und "Vorbeireden", wobei das Merkmal "maniriert" sogar einen Quotienten von 1 zwischen T0 und T28 aufweist, was also auf ein völliges Verschwinden des Symptoms bei den vorherigen Merkmalsträgern hindeutet.

Bei dem Gesamtscore und den Faktoren der AMDP-Skala zeigte sich insgesamt ein früheres Auftreten der Besserungstendenz als bei den Einzelmerkmalen. Der Gesamtscore, das "paranoide", das "halluzinatorische", das "katatone", das "manisch-schizoaffektive", das "Hostilitäts-" und das "somatisch depressive" Syndrom waren im Friedman-Test signifikant verändert, davon außer dem "katatonen" und dem "somatisch depressiven", alle auf dem 0,1 % Niveau. Der Gesamtscore, das "paranoide" und das "halluzinatorische" Syndrom veränderten sich im Dunn-Rankin-Test schon am 14. wie auch am 28. Tag auf dem 1 % Niveau; das "manisch-schizoaffektive" am 14. Tag auf

dem 5 % Niveau und am 28. Tag auf dem 1 % Niveau. Das "katatone" und das "Hostilitätssyndrom" veränderten sich beide erst am 28. Tag, das erste auf dem 5 % Niveau, das zweite bei 1 %. Die höchsten Mittelwerte wiesen (in der Reihenfolge der Höhe) das "paranoide", das "Hostilitäts-" und das "halluzinatorische" Syndrom auf.
Alle drei zeigten im Verlauf eine Rückbildung der absoluten Werte um mehr als 50 %, ebenso der Gesamtscore.
Von den 17 Items der BPRS-Skala erschienen in der Q-Statistik 5 signifikant verändert, gegenüber 10 im Friedman-Test; bei beiden erschien das Item "maniriert" auf dem 0,1 % Niveau verändert, außer diesem war nur noch das Item "Zerfall der Denkprozesse" in der Q-Statistik auf demselben Niveau signifikant. Im Friedman-Test erschienen zusätzlich die Items: "Angst", "Gespanntheit", "Größenerleben", "mangelnde Kooperation", "ungewöhnliche Denkinhalte".
Im Symmetrietest zeigten sich drei Items signifikant verbessert: "Zerfall der Denkprozesse" am 14. Tag und am 28. Tag, "Mißtrauen/paranoide Inhalte" am 14. Tag und "Maniriertheit" am 28. Tag.
Im Dunn-Rankin-Test erschienen fünf Items verändert, davon "Zerfall der Denkprozesse" ebenfalls am 14. und 28. Tag, ebenso wie "ungewöhnliche Denkinhalte" und "Mißtrauen/paranoide Inhalte"; "Angst" und "Maniriertheit" veränderten sich erst am 28. Tag.
Die höchsten Mittelwerte zeigten die Items "ungewöhnliche Denkinhalte", "Mißtrauen/paranoide Inhalte" und "Angst".
Bei den Verlaufsmaßen zeigte das Item "Maniriertheit" als einziges eine geringe Provokationstendenz, alle anderen hatten einen Provokationsquotienten von Null. Bei den Besserungsquotienten zeigte sich

"Zerfall der Denkprozesse" am 14. und am 28. Tag am stärksten gebessert, "Angst" und "ungewöhnliche Denkinhalte" am schwächsten, "Maniriertheit" erschien erst am 28. Tag deutlich besser.
Beim Gesamtscore und den Faktoren ließ sich ein ähnlicher Trend feststellen wie bei der AMDP-Skala; sie veränderten sich ebenfalls schon am 14. Tag, und zwar vor allem die Faktoren, die die höchsten Mittelwerte aufwiesen. Im Friedman-Test waren der Gesamtscore, die Faktoren "Denkstörungen", "mißtrauische Feindseligkeit", "ängstliche Depression" und die beiden Kontrastfunktionen signifikant verändert. Der Gesamtscore, "mißtrauische Feindseligkeit" und die Kontrastfunktion "coping versus resignation" veränderten sich im Dunn-Rankin-Test sowohl am 14. als auch am 28. Tag auf dem 1 % Niveau; "Denkstörungen" veränderte sich am 14. Tag auf dem 1 % Niveau und am 28. Tag auf dem 5 % Niveau, während sich "ängstliche Depression" und die Kontrastfunktion "schizophrenic versus depressiv" erst am 28. Tag änderten.
Vergleichend läßt sich zur Verlaufsbeschreibung anhand der AMDP- und BPRS-Skala sagen, daß auf Itemebene ein Unterschied besteht hinsichtlich der Anzahl der Items, die signifikant verändert erscheinen; zum zweiten bezüglich des Zeitpunktes der Veränderungen und drittens auch bezüglich der Höhe der Mittelwerte zum Zeitpunkt TO. Bei der AMDP-Skala erscheinen im Symmetrietest 10 Items signifikant verändert, davon 2 schon zum Zeitpunkt T14, im Dunn-Rankin 7, davon 1 zum Zeitpunkt T14. Dies erscheint, gemessen an der Gesamtzahl der Items, relativ wenig. Bei der BPRS erscheinen im Dunn-Rankin 5 Items signifikant verändert, davon 3 schon zum Zeitpunkt T14, im

Symmetrietest 3, davon 2 zum Zeitpunkt T14.

Betrachtet man die Items bezüglich ihres Inhaltes, so scheinen sowohl die signifikant veränderten Items der AMDP als auch der BPRS-Skala gleichbleibend eine Wahnsymptomatik abzufragen. Dabei ist auffällig, daß vor allem die Items der BPRS mit den höchsten Mittelwerten sich schon am 14. Tag verbessern, im Gegensatz zu den Items der AMDP, die einmal ebenfalls die höchsten Mittelwerte aufweisen, zum anderen ebenfalls eine Wahnsymptomatik abfragen, sich aber erst am 28. Tag signifikant verändern. Betrachtet man außerdem die Mittelwerte zum Zeitpunkt T0, so fällt weiter auf, daß, gemessen an der Skalierung, bei der AMDP von 0 bis 3 und bei der BPRS von 1 bis 7, die Mittelwerte bei der BPRS deutlich höher liegen. Dies alles deutet darauf hin, daß die Items der BPRS wesentlich inhaltsreicher sind und von daher eher als Syndrome aufzufassen sind und mit den Items der AMDP unmittelbar nicht zu vergleichen sind. Deutlich wird dies zum Beispiel an dem Item "maniriert" der AMDP-Skala, bzw. dem der BPRS "Maniriertheit", das sich nach dem Besserungsquotienten als Item der AMDP vollständig zurückbildet, als Item der BPRS nur zu ungefähr 50 %. Bei beiden Skalen änderte es sich allerdings gleichzeitig am 28. Tag.

Bezüglich des Gesamtscore und der Faktoren der beiden Skalen ist der beobachtete Verlauf sehr ähnlich. Der Gesamtscore und diejenigen Faktoren der beiden Skalen, die die höchsten Mittelwerte aufweisen, verändern sich alle schon zum Zeitpunkt T14 und zeigen auch weiterhin eine sehr gute Besserungstendenz. Es ist allerdings auch hier wieder auffällig, daß die Faktoren der BPRS einen relativ höheren

Mittelwert aufweisen, als diejenigen der AMDP.

Zusammenfassend lassen sich über die AMDP- und die BPRS-Skala hinsichtlich ihrer Brauchbarkeit als Verlaufsmeßinstrument folgendes sagen:

Die BPRS erscheint als eine relativ kurze Skala, die dadurch, daß die Items unmittelbar erläutert sind, gar kein, oder wenig Training erfordert. Ihre Items sind allerdings im Gegensatz zur AMDP-Skala wesentlich weniger eng formuliert, sodaß sie eigentlich eher den Charakter von Faktoren besitzen. Sowohl die Mittelwerte des Gesamtscore, der Faktoren als auch der Einzelitems erschienen, relativ zur Stufigkeit der Skala, in der BPRS insgesamt höher. Auch dies könnte als ein Hinweis dafür interpretiert werden, daß die Items der BPRS inhaltlich wesentlich weiter formuliert wurden als diejenigen der AMDP-Skala. In der statistischen Auswertung zeigten sich sowohl im Friedman-als auch im Dunn-Rankin-Test mehr Items signifikant verändert als im Cochran und McNemar, sodaß angenommen werden kann, daß ein Mehr an eingehender Information, insofern, als die Daten nicht dichotomisiert wurden, auch ein Mehr an ausgehender Information brachte. Die BPRS stellt sich also als relativ ökonomische Skala dar, die gut geeignet erscheint, Verläufe in globaler Art darzustellen. Bei der AMDP handelt es sich dagegen um eine wesentlich umfangreichere Skala, die ein umfangreiches Training erfordert und relativ zeitaufwendig ist. Die Items der AMDP weisen allerdings eine wesentlich größere Homogenität auf als diejenigen der BPRS. Dies zeigt sich unter anderem darin, daß nur sehr wenige, sehr spezifische Items einen relativ hohen Mittelwert aufweisen. In der statistischen Aus-

wertung zeigt sich eine Diskrepanz zwischen den erwarteten und den tatsächlichen Ergebnissen bezüglich der Lokationsteste McNemar und Dunn-Rankin insofern, als sich im McNemar, trotz der Dichotomisierung der Daten, mehr Items als signifikant verändert zeigen als im Dunn-Rankin-Test; und zwar tritt dies bei der BPRS stärker als bei der AMDP auf. Dies könnte einmal dahingehend erklärt werden, daß, wie bereits oben erwähnt, im Dunn-Rankin zur Lokationsprüfung alle Zeitpunkte gleichzeitig eingehen, während im McNemar ein Paarvergleich stattfindet, und zum anderen, daß bei der Dichotomisierung einer vierstufigen Skala ein geringerer Informationsverlust entsteht als bei einer siebenstufigen.

Bei beiden Skalen tritt eine Verbesserung im Verlauf zuerst im Gesamtscore und den Faktoren auf, während sich die Items in der Regel erst relativ spät ändern und zwar tritt dieser Trend bei der AMDP stärker auf, was dafür spricht, daß dabei die Homogenität eine große Rolle spielt: relativ homogene Items, bzw. Faktoren zeigen erst spät eine Veränderungstendenz.

Bezüglich der Interrater-Reliabilität schneiden beide Skalen gleich gut ab, wenn man davon absieht, daß ein relativ großer Anteil der Items der AMDP bezüglich des Kappakoeffizienten nicht interpretierbar sind. Auch dies ist wiederum auf die Homogenität der Items der AMDP-Skala zurückzuführen, insofern, als daß eine geringe Varianz des Merkmals, gleichgültig, ob nicht beobachtet oder nicht vorhanden, die Interpretation von Kappa unmöglich macht.

Zu den beiden Verfahren, dem Friedman-und Dunn-Rankin-Test, die bei einer derartigen Studie zum ersten Mal zur Anwendung kamen, sollte

noch bemerkt werden, daß sie auf der einen Seite rechnerisch einen größeren Aufwand darstellen als die Q-Statistik von Cochran und der McNemar-Symmetrietest, auf der anderen Seite bieten sie jedoch einige grundlegende Vorteile. Zum einen ist es bei diesen Verfahren methodisch unbedenklich, die Verlaufsänderung auch beim Gesamtscore und den Faktoren zu berechnen, die ja, wie sich in dieser Arbeit gezeigt hat, eher einen Hinweis auf eine Besserung liefern. Zum anderen ist es nicht sinnvoll, eine Rating-Skala durch die Dichotomisierung um einen Großteil ihrer Information zu berauben; vor allem, wenn es darauf ankommt, relativ feine Veränderungen festzustellen, was ja erst überhaupt durch die Stufigkeit einer Rating-Skala möglich ist.

Literaturverzeichnis:

Angst, J. et al: Das Dokumentationssystem der Arbeitsgemeinschaft für Methodik und Dokumentation in der Psychiatrie (AMDP)
Arzneim.-Forsch.19, 399 - 405 (1969)

Bartko, John J. et al: Selecting Techniques for evaluating the Reliability of psychiatric Data:
A Report from the international Pilot Study of Schizophrenia, from Mental Health Service System Reports: Multivariate statistical Methodologies used in the international Pilot Study of Schizophrenia,
Edited bei Bartko, John J. and Gulbinat, Walter;
DHHS Publication No. (ADM) 80 - 630, 1980

Baumann, Urs: Diagnostische Differenzierungsfähigkeit von Psychopathologie-Skalen, Arch. Psychiat. Nervenkr. 219, 89 - 103, 1974

Baumann, U. u. Angst, J.: Verlaufsanalysen psychopathologischer und somatischer Symptome
Arch. Psychiat. Nervenkr. 223, 227 - 238, 1977

Baumann, U. u. Woggon, B.: Interrater-Reliabilität bei Diagnosen, AMP-Syndromen und AMP-Symptomen
Arch. Psychiat. Nervenkr., 227, 3 - 15, 1979

Bente, D. et al: 2. Mitteilung: Verlaufsprofile psychopathologischer und somatischer Merkmale aus: Multidimensionale pharmakopsychiatrische Untersuchungen mit dem Neuroleptikum Perazin Pharmakopsychiat. 7, 8 - 17, 1974

Busch, H. u. Helmchen, H.: Dokumentation psychiatrischer Therapie Nervenarzt 44, 569 - 575, 1973

Cicchetti, Domenic V.: Assessing Inter-Rater Reliability for Rating Scales: Resolving some Basic Issues Brit. J. Psychiat. 129, 1976

Cohen, J.: Weighted Kappa: Nominal Scale Agreement with Provision for scaled Disagreement or partial Credit, Psychol. Bull., Vol. 70, No. 4 1968

Cohen, J.: A Coefficient of Agreement for Nominal Scales Educational and Psychological Measurement, 20, 37 - 46, 1960

Fleiss, J.L. et al: Large Sample Standard Errors of Kappa and Weighted Kappa, Psychol. Bull. Vol. 72 1969

Freudenthal, K. et al: AMP (PAS) and BPRS: A Comparison of two Assessment Methods, Pharmakopsychiat. 10 57 - 66, 1977

Gebhardt, R. u. Helmchen, H.: Zur Zuverlässigkeit psychopathologischer Symptomerfassung. Die Übereinstimmung mehrerer Beurteiler im AMP-System,

Schweizer Archiv für Neurologie, Neurochirurgie und Psychiatrie, Band 112, Heft 2 1973

Hadass, H. et al: 3. Mitteilung: Perazin-Konzentration im Blut und klinisches Wirkungsbild, aus: Multidimensionale pharmakopsychiatrische Untersuchungen mit dem Neuroleptikum Perazin Pharmakopsychiat. 7, 17 - 30, 1974

Hall, John N.: Inter-Rater Reliability of Ward Rating Scales Brit. J. Psychiat. 125, 248 - 255, 1974

Helmchen, H. u. Hippius, H.: 1. Mitteilung: Zielsetzung und methodische Grundlagen, aus: Multidimensionale pharmakopsychiatrische Untersuchungen mit dem Neuroleptikum Perazin, Pharmakopsychiat. 7, 1 - 8, 1974

Helmchen, H. et al: 7. Mitteilung: Methodologische Ergebnisse, aus: Multidimensionale pharmakopsychiatrische Untersuchungen mit dem Neuroleptikum Perazin, Pharmakopsychiat. 7, 58 - 64, 1974

Hentschel, U. et al: Diagnostische Klassifikationsversuche mit zwei standardisierten psychiatrischen Schätzskalen, Arch. Psychiat. Nervenkr. 221, 283 - 301, 1976

Lienert, G.A.: Verteilungsfreie Methoden in der Biostatistik Band 1, 191 - 200, 346 - 359, Band 2, 285 - 372 636 - 647, 1978

Maurer-Groeli, Y.A.: Untersuchung zur Interrater-Reliabilität des AMP-Systems, Arch. Psychiat. Nervenkr. 221, 321 - 329, 1976

Maxwell, A. E.: Coefficients of Agreement between Observers and their Interpretation, Brit. J. Psychiat. 130, 79 - 83, 1977

Mombour, W.: Syndrome bei psychiatrischen Erkrankungen Arch. Psychiat. Nervenkr. 219, 331 - 360 1974

Mombour, W.: Symptomhäufigkeiten bei psychiatrischen Erkrankungen, Arch. Psychiat. Nervenkr. 219, 133 - 152, 1974

Mombour, W. et al: Die Brief Psychiatric Rating Scale (BPRS) von Overall und Gorham bei akuten, paranoid-halluzinatorischen Psychosen. Untersuchungen zu einer deutschen Übersetzung der BPRS Pharmakopsychiat. 8, 279 - 288, 1975

Mombour, W. et al: Die Objektivierung psychiatrischer Syndrome durch multifaktorielle Analyse des psychopathologischen Befundes, Nervenarzt, 44, 352 - 358, 1973

Overall, J. E.: The Brief Psychiatric Rating Scale in Psychopharmacology Research, Pharmacopsychiat. Vol. 7, 67 - 78, 1974

Overall, J. E. u. Gorham, D. R.: The Brief Psychiatric Rating Scale Psychol. Rep. 10, 799 - 812, 1962

Overall, J. E. u. Hippius, H.: Psychiatric diagnostic Concepts among German-speaking Psychiatrists Comprehens. Psychiat. 15, 103 - 117, 1974

Pietzcker, A. et al: Ein Vergleich nosologisch-diagnostischer mit cluster-analytisch gefundenen Gruppen anhand AMP-dokumentierter psychopathologischer Befunde, Nervenarzt 48, 276 - 282, 1977

Skoda, C. et al: Prognostic Informativeness of the PSE and BPRS: A Comparison of their human and Computer Diagnostic Assessments, in: Multivariate Statistical Methodologies used in the International Pilot Study of Schizophrenia, edited by Bartko, J.J. and Gulbinat, W., DHHS Publication, No. (ADM) 80 - 630, 1980

Sulz-Blume, B. et al: Zur Stabilität der Faktorenstruktur der AMDP-Skala, Arch. Psychiat. Nervenkr. 227, 353 - 366, 1979

Woggon, B.: Einstufung von AMP-Symptomen bezüglich Fremd- und Selbstbeurteilung, Int. Pharmacopsychiat. 14, 158 - 169, 1979

Woggon, B.: Veränderungen der psychopathologischen Symptomatik während 20tägiger antidepressiver oder neuroleptischer Behandlung, psychiatrica clinica

Woggon, B. et al: Interrater-Reliabilität vom AMP-Symptomen Arch. Psychiat. Nervenkr. 225, 73 - 85, 1978

ABSTRACT

Holger Jung
Eine Untersuchung zum Vergleich zweier Einschätzskalen bei einer pharmakopsychiatrischen Studie mit akut schizophrenen Patienten.

In einer klinisch-pharmakologischen Studie wurden 28 akut schizophrene Patienten über eine Zeitdauer von 28 Tagen mit dem Neuroleptikum Perazin behandelt. Sie wurden dabei jeweils am Tag vor Beginn der Behandlung, am 14. und am 28. Tag der Behandlung einem Interview unterzogen, das, auf Video aufgenommen, anschließend von zwei an der AMDP-Skala trainierten Beurteilern auf der AMDP- und der BPRS-Skala eingeschätzt wurde. Anschließend wurde einmal die Interrater-Reliabilität mit Hilfe des Kappakoeffizienten von Cohen bestimmt, und zwar für die einzelnen Items als auch für Faktoren, die entsprechend früher durchgeführter Faktorenanalysen gebildet wurden. Zum anderen wurde ein Vergleich beider Skalen bezüglich ihrer Brauchbarkeit als Verlaufsinstrument angestellt. Dazu wurden beide Skalen einmal dichotomisiert und mit Hilfe der Q-Statistik von Cochran und dem McNemar-Test geprüft, zum anderen wurde das Datenmaterial polychotom belassen und dem Friedman-Test und dem d_{ij} Test von Dunn-Rankin unterzogen. In den Ergebnissen zeigte sich bezüglich der Interrater-Reliabilität, daß beide Skalen ähnlich gute Werte aufwiesen, daß aber nur ein Teil der Einzelitems der AMDP-Skala bezüglich des Kappa interpretiert werden konnte, bedingt durch die Homogenität der Stichprobe, die eine zu geringe Varianz von einigen Items bewirkte. Was die Vergleichbarkeit der beiden Skalen als Verlaufsinstrument betrifft, so ist festzustellen, daß sie auf Faktorenebene sehr ähnlich reagieren; auf Itemebene die BPRS jedoch früher Veränderungen aufweist, was auf die größere Komplexität ihrer Items hindeutet. In beiden Skalen verändern sich Items und Faktoren, die inhaltlich eine Wahnsymptomatik abfragen, am ehesten; sie zeigten auch insgesamt die höchsten Mittelwerte. Im direkten Vergleich stellt sich die BPRS gegenüber der AMDP als eine kurze, ökonomische Skala dar, deren Items einen wesentlich höheren Komplexitätsgrad aufweisen als diejenigen der AMDP. In dieser Studie erwiesen sich beide Skalen als ebenbürtig.

Lebenslauf

Am 2o.5.195o wurde ich, Holger Michael Jung, als Sohn des Kaufmanns Walter Jung und der Marlies Jung in Tuttlingen geboren. Im Frühjahr 1957 wurde ich in meinem Geburtsort eingeschult und trat dort nach 4 Jahren ins Gymnasium über. Im Juni 197o legte ich das Abitur ab und arbeitete anschliessend bis zu meiner Einberufung zum Wehrdienst am 1.1o.197o im Krankenhaus Tuttlingen als Hilfspfleger. Nach meiner Entlassung aus der Bundeswehr im März 1972 immatrikulierte ich mich an der Universität Tübingen zunächst im Fach Soziologie und im WS 72/73 dann im Fach Psychologie. Im SS 75, nachdem ich das Vordiplom im Fach Psychologie bestanden hatte, erhielt ich einen Studienplatz für das Fach Humanmedizin, ebenfalls an der Universität Tübingen. Nach bestandener ärztlicher Vorprüfung im WS 76/77 unterbrach ich mein Medizinstudium für 1 Jahr um die Diplomprüfung im Fach Psychologie abzulegen. Im WS 78/79 absolvierte ich dann den 1. Teil der ärztlichen Prüfung und im WS 8o/81 den 2. Teil. Als Wahlfach im Praktischen Jahr belegte ich das Fach Neurologie. Die ärztliche Prüfung legte ich am 21.4.1982 ab. Seit dem 1.8.82 arbeite ich als Assistenzarzt im Fach Anästhesie im Kreiskrankenhaus Kirchheim/Teck.